SIAM J. MATRIX ANAL. APPL.
Vol. 35, No. 3, pp. 886–903

ON GENERIC NONEXISTENCE OF THE SCHMIDT–ECKART–YOUNG DECOMPOSITION FOR COMPLEX TENSORS*

N. VANNIEUWENHOVEN†, J. NICAISE‡, R. VANDEBRIL†, AND K. MEERBERGEN†

Abstract. The Schmidt–Eckart–Young theorem for matrices states that the optimal rank-r approximation of a matrix is obtained by retaining the first r terms from the singular value decomposition of that matrix. This paper considers a generalization of this optimal truncation property to the rank decomposition (Candecomp/Parafac) of tensors and establishes a necessary orthogonality condition. We prove that this condition is not satisfied at least by an open set of positive Lebesgue measure in complex tensor spaces. It is proved, moreover, that for complex tensors of small rank this condition can be satisfied only by a set of tensors of Lebesgue measure zero. Finally, we demonstrate that generic tensors in cubic tensor spaces are not optimally truncatable.

Key words. canonical polyadic decomposition, Candecomp, Parafac, orthogonal rank decomposition, tensor singular value decomposition, Eckart–Young theorem

AMS subject classifications. 14A10, 14A25, 14Q15, 15A21, 15A69, 15B99

DOI. 10.1137/130926171

1. Introduction.

Data-sparse representation of elements living in the tensor product of finite dimensional vector spaces has become an intensely studied subject in recent years, yielding a myriad of factorizations, such as the higher-order singular value decomposition [17, 47, 48], the rank decomposition [26, 27] (also known as Candecomp/Parafac [10, 25] and CP decomposition), block-term decompositions [16], $\mathcal{H}$-Tucker [22, 24], and Tensor Trains [34], each with different assumptions and divergent applications. Among these, the rank decomposition is the oldest; according to [8], its roots, for symmetric tensors, can be traced back to algebraic geometry in the middle of 19th century as featured in the work of Sylvester. In contemporary terminology, a tensor $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2 \times \cdots \times n_d}$ with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ is said to admit a rank-R decomposition if it can be written as

$$\mathcal{A} = \sum_{i=1}^{R} \mathbf{a}_i^{(1)} \otimes \mathbf{a}_i^{(2)} \otimes \cdots \otimes \mathbf{a}_i^{(d)}, \tag{1.1}$$

*Received by the editors June 24, 2013; accepted for publication (in revised form) by P. Comon March 25, 2014; published electronically July 8, 2014. The second and third authors acknowledge the support by the Research Foundation–Flanders (FWO) project G034212N (Reestablishing Smoothness for Matrix Manifold Optimization via Resolution of Singularities).
http://www.siam.org/journals/simax/35-3/92617.html

†Numerical Approximation and Linear Algebra Group, Department of Computer Science, KU Leuven, 3001 Heverlee, Belgium (nick.vannieuwenhoven@cs.kuleuven.be, raf.vandebril@cs.kuleuven.be, karl.meerbergen@cs.kuleuven.be). The first author's research was supported by a Ph.D. fellowship of the Research Foundation–Flanders (FWO). The third and fourth authors acknowledge support by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization). The third author acknowledges the support of the KU Leuven Research Council, projects OT/11/055 (Spectral Properties of Perturbed Normal Matrices and their Applications) and CoE EF/05/006 Optimization in Engineering (OPTEC). The fourth author acknowledges the support of KU Leuven Research Council grants PFV/10/002 (Optimization in Engineering) and OT/10/038 (Multi-parameter Model Order Reduction and Its Applications).

‡Department of Mathematics, KU Leuven, 3001 Heverlee, Belgium (johannes.nicaise@wis.kuleuven.be).

where $\mathbf{a}_i^{(k)} \in \mathbb{F}^{n_k}$ and $\otimes$ denotes the Kronecker product, but not with fewer than R terms; it is the topic of this paper.

The rank decomposition is employed in a variety of scientific disciplines as a tool for data-driven analysis; much of the current interest in the rank decomposition originated in the psychometrics community from the works of Carroll and Chang [10] and Harshman [25]. This decomposition became known around 1980 in the field of chemometrics, where it is now well-entrenched [38]. Appellof and Davidson [2] employed it as an analytical technique for fluorescence spectroscopy and observed that the underlying physical process exactly admits a rank decomposition. Given the sampled tensor, its rank decomposition reveals the excitation and emission spectra for each of the chemical components in the fluorescent mixture. The review articles [30, 33] describe other applications within a data-driven setting.

The subject of this paper concerns a rank decomposition that mirrors the fundamental property of the singular value decomposition (SVD) of matrices: the approximation theorem of Schmidt [37, 44] and Eckart and Young [20]. Consider the SVD of $A \in \mathbb{F}^{m\times n}$ with $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$:

$$A = USV^T = \sum_{i=1}^{R} \sigma_i \mathbf{u}_i \otimes \mathbf{v}_i,$$

where $R \leq \min\{m, n\}$ is the rank of A, $S = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_R) \in \mathbb{R}^{R\times R}$, where the diagonal elements are assumed to be nonzero and sorted by decreasing magnitude, and U and V are matrices with orthonormal columns in the Euclidean inner product. We follow the convention from [23] using the transpose rather than the Hermitian conjugate, so that the above definition coincides with the definitions in the case of higher-order tensors. According to the Schmidt–Eckart–Young (SEY) theorem, the best rank-r approximation in the Euclidean topology is given by retaining the first r terms in the above sum. We are interested in a generalization of this theorem to the rank decomposition of tensors, i.e., we ask ourselves whether there exists a decomposition of $\mathcal{A} \in \mathbb{F}^{n_1\times n_2\times\cdots\times n_d}$ as in (1.1) such that retaining the first $r < R$ terms yields an optimal solution to the approximation problem

$$\sum_{i=1}^{r} \mathbf{a}_i^{(1)} \otimes \cdots \otimes \mathbf{a}_i^{(d)} \in \operatorname*{arg\,min}_{\mathbf{u}_j^{(k)} \in \mathbb{F}^{n_k}} \left\| \mathcal{A} - \sum_{i=1}^{r} \mathbf{u}_i^{(1)} \otimes \mathbf{u}_i^{(2)} \otimes \cdots \otimes \mathbf{u}_i^{(d)} \right\|, \tag{1.2}$$

where the norm is the Frobenius norm, also known as the Hilbert–Schmidt and ℓ_2 norm. We propose to call such a decomposition an SEY decomposition, and sometimes refer to it as an optimally truncatable decomposition.

The definition of the SEY decomposition is not vacuous: we will show that orthogonally diagonalizable tensors [13, 50] satisfy the above conditions. Such tensors appear in several applications and have been extensively studied [3, 9, 13, 36, 50]; nevertheless, it appears to be unknown that this decomposition is optimal in the above sense. In fact, we prove that orthogonal diagonalizability is not a necessary condition for an SEY decomposition, as we reveal a new class of optimally truncatable tensors.

There appears to be a consensus among researchers that an SEY decomposition is not feasible for tensors; however, general results on its (non)existence are scarce over the *real field* and lacking over the *complex field*. In this paper, we settle the case of complex tensors and prove that a generic tensor of small rank does not admit an SEY decomposition. The known results are covered in the next paragraphs.

Orthogonal tensor decompositions. Kolda [28, 29] investigated several orthogonal tensor decompositions as possible candidates for an SEY decomposition. In [29], she proved that an orthogonal tensor decomposition may not be optimal for a specific set of tensors over the real field, providing the first direct evidence that an SEY decomposition does not always exist. While [28] considers several orthogonal decompositions, it has, to our knowledge, never been proved that orthogonality is a necessary condition for optimal truncatability. We prove, in Theorem 3.3, that a form of orthogonality not considered before in the literature is necessary, while, in Theorem 3.6, we prove that another unconsidered form of orthogonality, generalizing results from [13, 28], is sufficient.

Ill-posedness. A popular argument for dismissing the existence of an SEY decomposition involves the ill-posedness of approximation problem (1.2), i.e., the solution may not exist. This problem arises because the set of rank-r tensors may not be closed, a result known to classical algebraic geometers, particularized by Bini, Lotti, and Romani [4, 5, 6], and recently scrutinized by de Silva and Lim [19]. This implies that some rank-r tensors can be approximated arbitrarily well by a tensor of rank $r_b < r$; the smallest of these ranks r_b is called the *border rank* of the tensor [6]. A tensor $\mathcal{A}$ whose border rank $r_b = \underline{\operatorname{rank}}_\otimes(\mathcal{A})$ differs from its rank $r = \operatorname{rank}_\otimes(\mathcal{A})$ will be referred to as an *open boundary tensor* (OBT). Approximation problem (1.2) is ill posed if all solutions of

$$\min_{\underline{\operatorname{rank}}_\otimes(\mathcal{B})\le r} \|\mathcal{A}-\mathcal{B}\|$$

are OBTs.[1] As an SEY decomposition can only exist if a solution exists, the following corollary is readily established.

Corollary 1.1. *A tensor $\mathcal{A} \in \mathbb{F}^{n_1\times\cdots\times n_d}$ does not admit an SEY decomposition if its best rank-r approximation does not exist for some r.*

Several specific examples of such tensors exist. The occurrence of OBTs was already exploited by Bini, Lotti, and Romani [4, 5, 6] in 1980 to derive original fast algorithms for approximate, but arbitrarily accurate, matrix multiplication. Another example is the rank-3 tensor $\mathbf{u}\otimes\mathbf{u}\otimes\mathbf{v}+\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{u}+\mathbf{v}\otimes\mathbf{u}\otimes\mathbf{u}$ with $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ linearly independent, which can be approximated arbitrarily well by a tensor of rank 2 [19]. Other specific examples of OBTs were investigated by Paatero [35]. In [40], Stegeman considered the occurrence of OBTs in some specific tensor spaces with two typical ranks.[2]

A more general result was obtained by Stegeman in [39, 41], in which it was shown that OBTs can occur in *real* tensor spaces of the form $2\times p\times q$ with positive volume when approximating a tensor of supergeneric[3] rank by one of the generic rank. Results for arbitrary third-order real tensor spaces were obtained by de Silva and Lim [19]. They proved that, in such a space, the set of tensors not admitting an optimal rank-2 approximation has positive measure [19, Theorem 8.4]; in other words, by selecting "random" tensors, there is a nonzero probability of obtaining one without a best rank-2 approximation. de Silva and Lim's theorem cannot be extended straightforwardly to complex tensors, as is clearly stated by its authors [19, section 9]. Nevertheless, as an immediate corollary of this theorem, one obtains the following result.

[1] Solving this optimization problem, which differs from (1.2), is difficult, as it requires insight into the limit points of sequences of rank-r tensors, which are currently not well understood. Some results in this direction have been established in [18, 19, 42, 43].

[2] Due to the lack of algebraic closedness of the real field, there may be multiple ranks occurring with a positive measure in the space [19, 31].

[3] The generic rank of a real or complex tensor space is, by definition, the smallest R such that the Zariski closure of the set of tensors of complex rank $\le R$ is the encompassing space $\mathbb{C}^{n_1\times\cdots\times n_d}$ [31, section 5.2.1]. A tensor of supergeneric, respectively, subgeneric, rank is one whose rank is larger, respectively, smaller, than the generic rank.

COROLLARY 1.2. *The set of third-order tensors in* $\mathbb{R}^{n_1 \times n_2 \times n_3}$ *that do not admit an SEY decomposition is nonempty and of positive Lebesgue measure.*

Beyond ill-posedness. The ill-posedness of approximation problem (1.2) appears to be cited often as the definite reason why an SEY decomposition cannot exist. For instance, the review article [30] states

> de Silva and Lim show, moreover, that the set of tensors of a given size that do not have a best rank-k approximation has positive volume (i.e., positive Lebesgue measure) for at least some values of k, so this problem of a lack of a best approximation is not a "rare" event.

The above statement is valid only for third-order real tensors; Corollary 1.2 cannot be generalized to higher orders or complex tensor spaces by using the techniques from [19, section 8] because they are founded on a classification of the finite number of orbits of the general linear group $\mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R}) \times \mathrm{GL}_2(\mathbb{R})$, which is, as the authors clearly state in [19, p. 1116],

> in general, not possible for tensors of arbitrary size and order simply because the dimension or "degrees of freedom" of $\mathbb{R}^{d_1 \times \cdots \times d_k}$ exceeds that of $\mathrm{GL}_{d_1,\ldots,d_k}(\mathbb{R})$ as soon as $d_1 \cdots d_k > d_1^2 + \cdots + d_k^2$ (which is almost always the case).

While the ill-posedness argument proves, for third-order real tensors, that a set of nonzero measure exists wherein no tensor admits an SEY decomposition, the general case is still open. All results discussed above rely on the existence of a set of positive measure consisting of the tensors admitting a supergeneric rank. This class of arguments fails for complex tensor spaces, as no such sets of positive measure exist.[4]

The main contribution of this paper is a proof that in every *complex* tensor space there exists a set of positive Lebesgue measure wherein no tensor admits an SEY decomposition, extending Corollary 1.2 to complex tensor spaces of arbitrary order.

THEOREM 1.3. *Let* $d \geq 3$. *The set of tensors in* $\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}$ *that do not admit an SEY decomposition is nonempty and of positive Lebesgue measure.*

Our second result states that for tensors of small rank, the subset consisting of tensors that do admit an SEY decomposition has measure zero.

THEOREM 1.4. *Let* $d \geq 3$. *The set of tensors of rank* r *in* $\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}$ *that admit an SEY decomposition has Lebesgue measure zero if the* r*th order secant variety of the Segre variety* $\mathbb{P}\mathbb{C}^{n_1} \times \mathbb{P}\mathbb{C}^{n_2} \times \cdots \times \mathbb{P}\mathbb{C}^{n_d}$ *is nondefective.*

Finally, we will even show that generic tensors in cubic tensor spaces do not admit an SEY decomposition.

THEOREM 1.5. *Let* $d \geq 3$, $(d, n) \neq (3, 3)$, *and* $(d, n) \neq (4, 2)$. *Then, a generic tensor in* $\mathbb{C}^{n \times n \times \cdots \times n}$ *(d times) does not admit an SEY decomposition.*

The outline of this paper is as follows. In the next section, some terminology and notation is fixed. Then, in section 3, the SEY decomposition is formally proposed. Both a necessary condition and a sufficient condition for its existence are investigated. In section 4, we will prove Theorems 1.3, 1.4, and 1.5. Finally, our conclusions are presented in section 5.

2. Preliminaries.

Notation. Throughout this paper, the symbol $\mathbb{F}$ denotes either the real field $\mathbb{R}$ or the complex field $\mathbb{C}$. Tensors are typeset in calligraphic uppercase letters ($\mathcal{A}$, $\mathcal{B}$), matrices in uppercase letters (A), vectors in boldface lowercase letters ($\mathbf{a}$, $\mathbf{u}$, $\mathbf{v}$), and

[4] The validity of this statement can be assessed by replacing the Veronese by the Segre variety in Corollary 6.11 of [15]; alternatively, see [31, p. 69].

scalars in lowercase letters (a, σ, λ). The scalar d is reserved for the order of a tensor. The Euclidean inner product is denoted by $\langle \cdot, \cdot \rangle$, the Euclidean norm by $\|\cdot\|$, and the tensor product by $\otimes$. The complex unity will be denoted by $\imath$.

Multilinear algebra. A tensor of order d is an element of the tensor product of d vector spaces: $\mathcal{A} \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \cdots \otimes \mathbb{F}^{n_d}$. Defining the standard tensor basis of order d as $\{\mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \cdots \otimes \mathbf{e}_{i_d}\}_{i_1,i_2,\ldots,i_d=1}^{n_1,n_2,\ldots,n_d}$, where $\mathbf{e}_{i_k}$ is the i_kth standard basis vector of $\mathbb{F}^{n_k}$, we can represent

$$\mathcal{A} = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} a_{i_1,i_2,\ldots,i_d} \mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \cdots \otimes \mathbf{e}_{i_d}$$

with respect to the standard tensor basis as the d-array

$$[a_{i_1,i_2,\ldots,i_d}]_{i_1,i_2,\ldots,i_d=1}^{n_1,n_2,\ldots,n_d} \in \mathbb{F}^{n_1 \times n_2 \times \cdots \times n_d}.$$

The Euclidean inner product of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{F}^{n_1 \times \cdots \times n_d}$ can then be defined as

$$\langle \mathcal{A}, \mathcal{B} \rangle := \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} a_{i_1,\ldots,i_d} \overline{b}_{i_1,\ldots,i_d}$$

with $\overline{x}$ the complex conjugate of x. The corresponding Euclidean norm is $\|\mathcal{A}\| := \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$. Note that this definition coincides with the Euclidean norm of $\mathcal{A}$ when it is considered as an element of $\mathbb{F}^{n_1 \cdots n_d}$. We call $\mathcal{A}$ and $\mathcal{B}$ *simple* or *rank-1* tensors if

$$\mathcal{A} = \mathbf{a}^{(1)} \otimes \mathbf{a}^{(2)} \otimes \cdots \otimes \mathbf{a}^{(d)} \text{ and } \mathcal{B} = \mathbf{b}^{(1)} \otimes \mathbf{b}^{(2)} \otimes \cdots \otimes \mathbf{b}^{(d)} \text{ with } \mathbf{a}^{(k)}, \mathbf{b}^{(k)} \in \mathbb{F}^{n_k}.$$

In this case, the inner product simplifies to

$$\langle \mathcal{A}, \mathcal{B} \rangle = \langle \mathbf{a}^{(1)}, \mathbf{b}^{(1)} \rangle \cdots \langle \mathbf{a}^{(d)}, \mathbf{b}^{(d)} \rangle;$$

see, e.g., [23, section 4.5.1]. Two tensors $\mathcal{A}$ and $\mathcal{B}$ are orthogonal, $\mathcal{A} \perp \mathcal{B}$, iff $\langle \mathcal{A}, \mathcal{B} \rangle = 0$. Two simple tensors $\mathcal{A}$ and $\mathcal{B}$ are thus orthogonal iff there is at least one mode k wherein the mode-k vectors are orthogonal, $\mathbf{a}^{(k)} \perp \mathbf{b}^{(k)}$.

Algebraic geometry. A tensor $\mathcal{A}$ is of *rank* r if it can be written as

$$\mathcal{A} = \sum_{i=1}^{r} \lambda_i \mathbf{a}_i^{(1)} \otimes \mathbf{a}_i^{(2)} \otimes \cdots \otimes \mathbf{a}_i^{(d)} \quad \text{with} \quad \mathbf{a}_i^{(k)} \in \mathbb{F}^{n_k} \quad \text{and} \quad \lambda_i \in \mathbb{F},$$

but not with fewer than r terms. This rank will be denoted by $\operatorname{rank}_{\otimes}(\mathcal{A})$.

It is often useful to consider tensors up to scalar multiplication, i.e., as $\mathbb{F}$-rational points in the projective space

$$P = \mathbb{P}(\mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \cdots \otimes \mathbb{F}^{n_d}).$$

Then, simple tensors correspond precisely to the $\mathbb{F}$-rational points of the Segre variety

$$\mathcal{S}_{n_1,n_2,\ldots,n_d}^{\mathbb{F}} = \mathbb{P}\mathbb{F}^{n_1} \times \mathbb{P}\mathbb{F}^{n_2} \times \cdots \times \mathbb{P}\mathbb{F}^{n_d} \subset P.$$

If $\mathbb{F} = \mathbb{C}$, the tensors of border rank at most r are described by points on the rth *secant variety* $\sigma_r(\mathcal{S}_{n_1,n_2,\ldots,n_d}^{\mathbb{C}})$ of the Segre variety $\mathcal{S}_{n_1,n_2,\ldots,n_d}^{\mathbb{C}}$ [31]. Recall, from, e.g., [31], that the secant variety of a subvariety $\mathcal{S}$ of a projective space $\mathbb{P}\mathbb{C}^N$ is defined as the Zariski closure of the union of the linear span of r points on the variety $\mathcal{S}$:

$$\sigma_r(\mathcal{S}) := \overline{\bigcup_{p_1,\ldots,p_r\in\mathcal{S}} \operatorname{span}(p_1, p_2, \ldots, p_r)}.$$

A $\mathbb{C}$-rational point $\mathcal{A}$ of P is of border rank r if it belongs to $\sigma_r(\mathcal{S}^{\mathbb{C}}_{n_1,n_2,\ldots,n_d})$ but not to $\sigma_{r-1}(\mathcal{S}^{\mathbb{C}}_{n_1,n_2,\ldots,n_d})$. By construction, the tensors of rank r form a Zariski-dense constructible subset of the rth secant variety: rank-r tensors are generic within the set of tensors of border rank r.

For more information about the connection between algebraic geometry and multilinear algebra, see Landsberg's book [31].

3. SEY decomposition. We consider a natural generalization of the SEY theorem to higher-order tensors.

DEFINITION 3.1. *A tensor $\mathcal{A} \in \mathbb{F}^{n_1\times n_2\times\cdots\times n_d}$ admits an SEY decomposition iff it can be written as a linear combination of simple tensors,*

$$\mathcal{A} = \sum_{i=1}^{R} \sigma_i \mathcal{A}_i = \sum_{i=1}^{R} \sigma_i \mathbf{a}_i^{(1)} \otimes \mathbf{a}_i^{(2)} \otimes \cdots \otimes \mathbf{a}_i^{(d)}$$

with $R = \operatorname{rank}_\otimes(\mathcal{A})$, $\sigma_i \in \mathbb{F}$, $\mathbf{a}_i^{(k)} \in \mathbb{F}^{n_k}$, and $\|\mathbf{a}_i^{(k)}\| = 1$, and such that truncating the decomposition is optimal for all $r = 1, 2, \ldots, R$:

$$\sum_{i=1}^{r} \sigma_i \mathcal{A}_i \in \operatorname*{arg\,min}_{\operatorname{rank}_\otimes(\mathcal{B})=r} \|\mathcal{A} - \mathcal{B}\|.$$

Note that there may be multiple optima for any given rank. We say that a point in $\mathbb{P}\mathbb{F}^{n_1\times n_2\times\cdots\times n_d}$ has an SEY decomposition if it has a tuple of homogeneous coordinates in $\mathbb{F}^{n_1\times n_2\times\cdots\times n_d}$ with an SEY decomposition.

The definition imposes no orthogonality constraints, and the σ_i's are not sorted by decreasing magnitude. It will, nonetheless, be shown in Corollary 3.4 that the SVD is the only SEY decomposition for second-order tensors.

3.1. A necessary condition. We begin by establishing a necessary condition for admitting an SEY decomposition, which we claim to be weak two-orthogonal.

DEFINITION 3.2. *A* weak two-orthogonal decomposition *is an orthogonal decomposition of an order-d tensor $\mathcal{A} \in \mathbb{F}^{n_1\times n_2\times\cdots\times n_d}$ in simple tensors,*

$$\mathcal{A} = \sum_{i=1}^{R} \sigma_i \mathbf{u}_i^{(1)} \otimes \mathbf{u}_i^{(2)} \otimes \ldots \otimes \mathbf{u}_i^{(d)} \quad \textit{with} \quad \mathbf{u}_i^{(k)} \in \mathbb{F}^{n_k},\ \|\mathbf{u}_i^{(k)}\| = 1,\ \textit{and}\ \sigma_i \in \mathbb{F},$$

and all terms are pairwise orthogonal in at least two modes:

$$\forall\, 1 \le i < j \le R : \exists\, 1 \le k_1 < k_2 \le d : \mathbf{u}_i^{(k_1)} \perp \mathbf{u}_j^{(k_1)} \quad \textit{and} \quad \mathbf{u}_i^{(k_2)} \perp \mathbf{u}_j^{(k_2)}.$$

Note that k_1 and k_2, in the above definition, depend on i and j: we could have written this more explicitly as $k_1(i,j) < k_2(i,j)$; however, this would have unnecessarily encumbered the notation.

THEOREM 3.3. *If $\mathcal{A}$ admits an SEY decomposition as in Definition* 3.1*, then it is a weak two-orthogonal decomposition.*

Proof. We prove the assertion by contradiction. Consider the following complete order on elements of $\mathbb{N} \times \mathbb{N}$:

$$(i_1, j_1) < (i_2, j_2) \qquad \text{iff} \qquad (i_1 < i_2) \text{ or } ((i_1 = i_2) \text{ and } (j_1 < j_2)).$$

Assume that an SEY decomposition is not weak two-orthogonal. Then, there exists a maximal tuple $(i^*, j^*) \in \{1, \ldots, R\} \times \{1, \ldots, R\}$ with respect to the above order, such that $\mathcal{A}_{i^*}$ and $\mathcal{A}_{j^*}$ are not two-orthogonal. Due to the symmetry of the inner product, we may assume that $i^* < j^*$. We distinguish between two cases: either $\mathcal{A}_{i^*}$ and $\mathcal{A}_{j^*}$ are orthogonal in only one mode, or they are not orthogonal.

Consider first the case in which $\mathcal{A}_{i^*}$ and $\mathcal{A}_{j^*}$ are not orthogonal. Let $1 \le k \le d$ be any mode, and consider the simple tensor

$$\epsilon \mathcal{P}_{i^*j^*}^{(k)} := \mathbf{a}_{i^*}^{(1)} \otimes \cdots \otimes \mathbf{a}_{i^*}^{(k-1)} \otimes \epsilon \mathbf{a}_{j^*}^{(k)} \otimes \mathbf{a}_{i^*}^{(k+1)} \otimes \cdots \otimes \mathbf{a}_{i^*}^{(d)} \text{ with } \epsilon \in \mathbb{R},$$

which has the property that it only perturbs the kth mode of the i^*th term in the SEY decomposition of $\mathcal{A}$ in the direction of $\mathbf{a}_{j^*}^{(k)}$. That is,

$$\mathbf{a}_{i^*}^{(1)} \otimes \cdots \otimes \mathbf{a}_{i^*}^{(d)} + \epsilon \mathcal{P}_{i^*j^*}^{(k)} = \mathbf{a}_{i^*}^{(1)} \otimes \cdots \otimes \mathbf{a}_{i^*}^{(k-1)} \otimes \left(\mathbf{a}_{i^*}^{(k)} + \epsilon \mathbf{a}_{j^*}^{(k)} \right) \otimes \mathbf{a}_{i^*}^{(k+1)} \otimes \cdots \otimes \mathbf{a}_{i^*}^{(d)}.$$

Consequently, adding this perturbation to $\sum_{j=1}^{j^*-1} \sigma_j \mathcal{A}_j$ does not increase its rank. Because $\mathcal{A}$ admits an SEY decomposition, this perturbed sum does not improve the approximation error:

$$\left\| \mathcal{A} - \sum_{j=1}^{j^*-1} \sigma_j \mathcal{A}_j - \epsilon \mathcal{P}_{i^*j^*}^{(k)} \right\|^2 \ge \left\| \mathcal{A} - \sum_{j=1}^{j^*-1} \sigma_j \mathcal{A}_j \right\|^2 = \eta^2. \tag{3.1}$$

On the other hand, by expanding the norm on the left-hand side and observing that the norm of $\epsilon \mathcal{P}_{i^*j^*}^{(k)}$ is ϵ^2, we obtain

$$\eta^2 - 2\Re \left\langle \mathcal{A} - \sum_{j=1}^{j^*-1} \sigma_j \mathcal{A}_j, \epsilon \mathcal{P}_{i^*j^*}^{(k)} \right\rangle + \epsilon^2 \ge \eta^2,$$

which after exploiting linearity and reordering the terms becomes

$$\epsilon^2 \ge 2\Re \left(\sum_{j=j^*}^{R} \sigma_j \langle \mathcal{A}_j, \epsilon \mathcal{P}_{i^*j^*}^{(k)} \rangle \right) = 2\Re \left(\epsilon \sum_{j=j^*}^{R} \sigma_j \langle \mathbf{a}_j^{(k)}, \mathbf{a}_{j^*}^{(k)} \rangle \prod_{\substack{m=1 \\ m \ne k}}^{d} \langle \mathbf{a}_j^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle \right), \tag{3.2}$$

where $\Re(\alpha)$ denotes the real part of α. From the maximality of (i^*, j^*), it follows that if $j > j^*$, then $\mathcal{A}_j$ is orthogonal to $\mathcal{A}_{i^*}$ in at least two modes. Otherwise, there would be a $j' > j^*$ such that $\mathcal{A}_{j'}$ is not two-orthogonal to $\mathcal{A}_{i^*}$, yielding an immediate contradiction to the maximality of (i^*, j^*). Consequently, $\prod_{m=1, m \ne k}^{d} \langle \mathbf{a}_j^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle = 0$ if $j > j^*$, because only one mode is excluded from this product, so that $\mathcal{A}_{i^*}$ is still orthogonal in at least one other mode to $\mathcal{A}_j$. Therefore, (3.2) simplifies to

$$\epsilon^2 \ge 2\epsilon \cdot \Re \left(\sigma_{j^*} \prod_{\substack{m=1 \\ m \ne k}}^{d} \langle \mathbf{a}_{j^*}^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle \right). \tag{3.3}$$

If $\mathbb{F} = \mathbb{C}$, an additional inequality is required to construct our contradiction; to this end, consider (3.1) again, but now apply the permutation $\imath \epsilon \mathcal{P}_{i^*j^*}^{(k)}$ with $\epsilon \in \mathbb{R}$ and expand the norm to obtain (3.2), with the only difference being that ϵ should now be replaced with $-\imath \epsilon$ on the right-hand side. Applying the induction hypothesis, as before, we find that

$$\epsilon^2 \geq 2\epsilon \cdot \Im\left(\sigma_{j^*} \prod_{\substack{m=1\\m\neq k}}^{d} \langle \mathbf{a}_{j^*}^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle\right), \tag{3.4}$$

where we have used the identity $\Re(-\imath\alpha) = \Im(\alpha)$, and $\Im(\alpha)$ denotes the imaginary part of α. If $\mathbb{F} = \mathbb{R}$, (3.4) is not required. If the real part and imaginary part on the right-hand sides of (3.3) and (3.4), respectively, are nonzero, a contradiction to the optimality of the SEY decomposition can be constructed by choosing ϵ sufficiently small. Therefore,

$$\sigma_{j^*} \prod_{\substack{m=1\\m\neq k}}^{d} \langle \mathbf{a}_{j^*}^{(m)}, \mathbf{a}_{i^*}^{(m)} \rangle = 0.$$

As $\sigma_{j^*} \neq 0$, because that would otherwise contradict $R = \operatorname{rank}_\otimes(\mathcal{A})$, there should be at least one $m \neq k$ such that $\mathbf{a}_{j^*}^{(m)} \perp \mathbf{a}_{i^*}^{(m)}$.

Consider now the second case; $\mathcal{A}_{i^*}$ and $\mathcal{A}_{j^*}$ are orthogonal in one mode, say, k. Repeat the argument for the previous case for mode k, which is now fixed. Then, we find that this case is also contradictory. Consequently, the assumption of the existence of such a maximal tuple (i^*, j^*) must be false. This concludes the proof. ∎

As an immediate consequence, we obtain the following corollary.

Corollary 3.4. *The SVD is the unique SEY decomposition of a second-order tensor $\mathcal{A} \in \mathbb{F}^{n_1 \times n_2}$.*

Proof. In the case of two modes, the condition in Theorem 3.3 simplifies to

$$\mathcal{A} = USV^T \quad \text{with } U \in \mathbb{F}^{n_1 \times R}, V \in \mathbb{F}^{n_2 \times R} \text{ and } S \in \mathbb{F}^{R \times R}, \tag{3.5}$$

where U and V have orthonormal columns, and S is a diagonal matrix. The diagonal of S can be chosen real and nonnegative, and we can assume it is sorted by decreasing magnitude. Thus, (3.5) is the compact SVD. ∎

Weak two-orthogonality is not a sufficient condition, because it does not exclude the OBT $\mathbf{u} \otimes \mathbf{u} \otimes \mathbf{v} + \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{u} + \mathbf{v} \otimes \mathbf{u} \otimes \mathbf{u}$ with $\mathbf{u}, \mathbf{v} \in \mathbb{F}^n$, and $\mathbf{u} \perp \mathbf{v}$ [19].

3.2. A sufficient condition. The Tensor SVD, or completely orthogonal rank decomposition [13, 28, 50], appears in several applications related to blind-source separation [3, 9, 36]. We show that such a decomposition is a special case of the strong two-orthogonal decomposition, which is proved to be an SEY decomposition.

Definition 3.5. *A rank-R tensor $\mathcal{A} \in \mathbb{F}^{n_1 \times \cdots \times n_d}$ admits a* strong two-orthogonal decomposition *of rank R with* splitting point $s \in \{1, 2, \ldots, d\}$ *if it can be written as*

$$\mathcal{A} = \sum_{i=1}^{R} \sigma_i \underbrace{\mathbf{u}_i^{(1)} \otimes \cdots \otimes \mathbf{u}_i^{(s)}}_{\mathcal{U}_i^s} \otimes \underbrace{\mathbf{v}_i^{(s+1)} \otimes \cdots \otimes \mathbf{v}_i^{(d)}}_{\mathcal{V}_i^s} = \sum_{i=1}^{R} \sigma_i \mathcal{U}_i^s \otimes \mathcal{V}_i^s$$

with $\mathbf{u}_i^{(k)} \in \mathbb{F}^{n_k}$, $\mathbf{v}_i^{(k)} \in \mathbb{F}^{n_k}$, $\sigma_i \in \mathbb{R}$, $\|\mathbf{u}_i^{(k)}\| = 1$, $\|\mathbf{v}_i^{(k)}\| = 1$, and

$$\forall\, 1 \leq i < j \leq R: \quad \mathcal{U}_i^s \perp \mathcal{U}_j^s \quad \text{and} \quad \mathcal{V}_i^s \perp \mathcal{V}_j^s.$$

The coefficients σ_i are assumed to be sorted: $\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_R^2 > 0$.

The fact that the σ_i's are required to be real does not limit generality; the coefficients can always be chosen to be the norm of $\sigma_i \mathcal{U}_i^s \otimes \mathcal{V}_i^s$, which is real. The partitioning of the modes is also not restricted to consecutive modes because we may arbitrarily reorder them; however, for simplicity of presentation, we assume a partitioning with consecutive modes in the remainder.

From the above definition, it follows that a strong two-orthogonal decomposition is a weak two-orthogonal decomposition with an additional restriction on the choice of k_1 and k_2 in Definition 3.2; it is, in addition, required that $1 \leq k_1 \leq s < k_2 \leq d$ for all combinations of two terms. This turns out to be sufficient for obtaining optimality.

THEOREM 3.6. *A rank-R strong two-orthogonal decomposition of a rank-R tensor is an SEY decomposition.*

Proof. A tensor space $\mathbb{T} := \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_s} \otimes \mathbb{F}^{n_{s+1}} \otimes \cdots \otimes \mathbb{F}^{n_d}$ is isomorphic, as a vector space, thus ignoring the tensor structure [23], to $\mathbb{T}_s := \mathbb{F}^{n_1 \cdots n_s} \otimes \mathbb{F}^{n_{s+1} \cdots n_d}$. By definition, a simple tensor $\mathbf{a}^{(1)} \otimes \cdots \otimes \mathbf{a}^{(s)} \otimes \mathbf{a}^{(s+1)} \otimes \cdots \otimes \mathbf{a}^{(d)} \in \mathbb{T}$ becomes the rank-1 matrix

$$(\mathbf{a}^{(1)} \otimes \cdots \otimes \mathbf{a}^{(s)}) \otimes (\mathbf{a}^{(s+1)} \otimes \cdots \otimes \mathbf{a}^{(d)}) \in \mathbb{T}_s \cong \mathbb{T},$$

where the products inside the brackets can be interpreted as Kronecker products. Using multilinearity, it follows that a tensor $\mathcal{A}$ satisfying Definition 3.5 with splitting point s admits the decomposition

$$\mathcal{A} = \sum_{i=1}^{R} \sigma_i \mathcal{U}_i^s \otimes \mathcal{V}_i^s,$$

which is, upon closer inspection, the compact SVD of $\mathcal{A}$ considered as matrix in $\mathbb{T}_s$. Observe that any rank-r decomposition in $\mathbb{T}$ is a matrix of rank at most r in $\mathbb{T}_s$. As the SVD provides an optimal approximation of rank r for the matrix $\mathcal{A} \in \mathbb{T}_s$, it follows that no rank-r decomposition over $\mathbb{T}$ can be a strictly better approximation than the provided strong two-orthogonal decomposition; otherwise, the optimality of the matrix SVD would be contradicted. Finally, considering a limit of a sum of r simple tensors over $\mathbb{T}$ cannot improve the SVD, because $\mathcal{A} \in \mathbb{T}_s$, which is an order-2 tensor product for which it is known that the set of rank-r tensors is closed; hence, limits do not extend the set over which the optimization is defined. $\square$

The Tensor SVD is a rank decomposition where orthogonality is imposed in every mode; hence, it is an SEY decomposition.

It remains an open question whether strong two-orthogonality is also necessary.

4. Generic nonexistence. In this section, the prime result is presented, which states that a set of nonzero Lebesgue measure exists in $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}$, $d \geq 3$, such that its elements do *not* exhibit an SEY decomposition. Note that we restrict ourselves to the complex case $\mathbb{F} = \mathbb{C}$ in this section; however, recall from the introduction that this is the interesting case as no general results are known. The results in this section apply for $d \geq 3$; matrices admit an SEY decomposition as per Corollary 3.4.

Throughout this section, we consider secants of the Segre variety $\mathcal{S}^{\mathbb{C}}_{n_1,\ldots,n_d}$ embedded in $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}$; henceforth, we let $\mathcal{S}$ denote $\mathcal{S}^{\mathbb{C}}_{n_1,\ldots,n_d}$. The set of tensors that admit a rank-r weak two-orthogonal decomposition, written $\sigma_r^{\perp}$, is a subset of the set of tensors of rank at most r, written σ_r', which, in turn, is a subset of the tensors of border rank at most r, i.e., the rth order secant variety $\sigma_r(\mathcal{S})$ of the Segre variety $\mathcal{S}$; henceforth, the former variety will be denoted by σ_r. Recall from section 2 that a generic tensor of border rank r also has rank r over $\mathbb{C}$.

Before proceeding with our proof strategy, we relate an interesting alternative strategy communicated to us by G. Ottaviani. From [14], we know that generic rank-r tensors in $\mathbb{P}\mathbb{C}^{n_1 \times n_2 \times \cdots \times n_d}$ are identifiable, at least if r is sufficiently small. This entails that the points on the Segre variety $\mathcal{S}^{\mathbb{C}}_{n_1,\ldots,n_d}$ are uniquely determined and in generic configuration. However, it is intuitively clear that the necessary condition we derived in section 3.1 imposes a certain configuration on the points, hereby contradicting genericity. This immediately entails that an SEY decomposition does not exist for such generic rank-r tensors. The strategy we follow in this paper can result in stronger statements, as we are able to use the conditions for nondefectivity of Segre varieties, rather than the more restrictive conditions for identifiability. In particular, we can demonstrate that a generic tensor in a cubic tensor space $\mathbb{C}^{n\times\cdots\times n}$, d times, with n sufficiently large, does not admit an SEY decomposition, a result that cannot be derived from the strategy sketched in this paragraph. It is an important observation worth stressing, however: *optimal truncation and identifiability cannot occur simultaneously in generic rank-r tensors.*

As a first remark, note that $\sigma_r^\perp$ is not a complex algebraic variety because of the complex conjugation that appears in the definition of the Euclidean inner product. Therefore, we propose to investigate the underlying *real algebraic structure* of $\sigma_r^\perp$ by applying a Weil restriction of scalars from $\mathbb{C}$ to $\mathbb{R}$; this simply means that we write each complex coordinate x as $x = u + \imath v$ and consider u and v as new coordinates over $\mathbb{R}$.[5] Let $\prod_{\mathbb{C}/\mathbb{R}}$ denote the Weil restriction functor, and then $\sigma_r^\perp$ can be considered as a Zariski-closed subset of the real algebraic manifold

$$\left(\prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^{n_1\times n_2\times\cdots\times n_d}\right)(\mathbb{R}) = \mathbb{PC}^{n_1\times n_2\times\cdots\times n_d};$$

here $A(\mathbb{R})$ denotes the set of real points of the real algebraic variety A. We will compare the size of σ_r and $\sigma_r^\perp$ by considering their dimensions as real algebraic manifolds.

Proposition 4.1. *We have*

$$\dim(\sigma_r^\perp) \le 2r\sum_{\ell=1}^{d}(n_\ell - 1) + 2\delta - 2$$

with $\delta = r - 2\lfloor r/2\rfloor$.

Proof. We will need an adapted version of the join operation. Let $V_1,\ldots,V_s$ be subvarieties of $\prod_{\mathbb{C}/\mathbb{R}}\mathbb{PC}^N$ for some positive integer N. Applying the Weil restriction functor to the projection morphism

$$\mathbb{A}_{\mathbb{C}}^N \setminus \{(0,\ldots,0)\} \to \mathbb{PC}^N,$$

where $\mathbb{A}_{\mathbb{C}}^N$ denotes the N-dimensional affine space over $\mathbb{C}$, yields a morphism of $\mathbb{R}$-varieties

$$\pi : \prod_{\mathbb{C}/\mathbb{R}}(\mathbb{A}_{\mathbb{C}}^N \setminus \{(0,\ldots,0)\}) \to \prod_{\mathbb{C}/\mathbb{R}}\mathbb{PC}^N$$

whose source is canonically isomorphic to $\mathbb{A}_{\mathbb{R}}^{2N}\setminus\{(0,\ldots,0)\}$. We define the affine cone of V_i as

$$\widetilde{V}_i := \pi^{-1}(V_i) \cup \{(0,\ldots,0)\}$$

[5] Alternatively, one could have attempted to investigate the Zariski closure of $\sigma_r^\perp$, considered as a complex algebraic variety; however, as an anonymous referee also remarked, "the complex Zariski closure . . . may be too large. Indeed, the complex Zariski closure of the subset of pairs of complex vectors $(\mathbf{x},\mathbf{y}) \in \mathbb{C}^n\times\mathbb{C}^n$, such that their Hermitian product vanishes, is [already] the whole Cartesian product $\mathbb{C}^n\times\mathbb{C}^n$." It seems that this strategy is not viable.

for each i in $\{1,\ldots,s\}$. This is a closed subvariety of $\mathbb{A}^{2N}_{\mathbb{R}}$. Now the Zariski closure of the image of $\widetilde{V}_1 \times_{\mathbb{R}} \cdots \times_{\mathbb{R}} \widetilde{V}_s$ under the addition morphism

$$(\mathbb{A}^{2N}_{\mathbb{R}})^s \to \mathbb{A}^{2N}_{\mathbb{R}} : (v_1,\ldots,v_s) \mapsto v_1 + \cdots + v_s$$

is a union of the origin and fibers of the projection morphism π. Removing the origin and taking the image under π, we obtain a closed subvariety of $\prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^N$ that we call the join of $V_1,\ldots,V_s$ and denote by $J(V_1,\ldots,V_s)$. The fibers of π are punctured real planes, so that the dimension of $J(V_1,\ldots,V_s)$ is at most

$$\dim(V_1) + \cdots + \dim(V_s) + 2(s-1).$$

The simple tensors in $\mathbb{PC}^{n_1} \times_{\mathbb{C}} \cdots \times_{\mathbb{C}} \mathbb{PC}^{n_d}$ with complex homogeneous coordinates

$$\begin{bmatrix} u_1^{(1)} + \imath v_1^{(1)} & \cdots & u_{n_1}^{(1)} + \imath v_{n_1}^{(1)} \end{bmatrix}^T \otimes \cdots \otimes \begin{bmatrix} u_1^{(d)} + \imath v_1^{(d)} & \cdots & u_{n_d}^{(d)} + \imath v_{n_d}^{(d)} \end{bmatrix}^T$$

correspond canonically to the real points of the $\mathbb{R}$-variety

$$V = \prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^{n_1} \times_{\mathbb{R}} \cdots \times_{\mathbb{R}} \prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^{n_d}.$$

For all i,j in $\{1,\ldots,d\}$ with $i \neq j$, the pairs of simple tensors that are orthogonal in modes i and j correspond to the real points of the subvariety W_{ij} of $V \times_{\mathbb{R}} V$ defined by the equations

$$\begin{cases} \sum_{\ell=1}^{n_i} (u_\ell^{(i)} \dot{u}_\ell^{(i)} + v_\ell^{(i)} \dot{v}_\ell^{(i)}) = 0, \\ \sum_{\ell=1}^{n_i} (u_\ell^{(i)} \dot{v}_\ell^{(i)} - \dot{u}_\ell^{(i)} v_\ell^{(i)}) = 0 \end{cases} \quad \text{and} \quad \begin{cases} \sum_{\ell=1}^{n_j} (u_\ell^{(j)} \dot{u}_\ell^{(j)} + v_\ell^{(j)} \dot{v}_\ell^{(j)}) = 0, \\ \sum_{\ell=1}^{n_j} (u_\ell^{(j)} \dot{v}_\ell^{(j)} - \dot{u}_\ell^{(j)} v_\ell^{(j)}) = 0, \end{cases}$$

where we used the coordinates $\dot{u}$ and $\dot{v}$ for points on the second factor of the product $V \times_{\mathbb{R}} V$. The variety $V \times_{\mathbb{R}} V$ is irreducible of dimension $4(n_1-1) + \cdots + 4(n_d-1)$, and we claim that W_{ij} has codimension 4. This is easy, but somewhat tedious, to check by covering the projective spaces $\mathbb{PC}^{n_q}$, $q = 1,\ldots,d$, by their standard affine charts, which gives rise to an open covering of $V \times_{\mathbb{R}} V$ by affine spaces.

We define W as the union of all the varieties W_{ij} with $i \neq j$ inside $V \times_{\mathbb{R}} V$. This is again a variety of dimension $4(n_1-1) + \cdots + 4(n_d-1) - 4$, whose $\mathbb{R}$-rational points correspond to pairs of simple tensors with complex homogeneous coordinates that are orthogonal in at least two modes.

Set $N = n_1 \cdots n_d$. Applying the Weil restriction functor to the Segre embedding, we get a closed immersion of $\mathbb{R}$-varieties

$$V \times_{\mathbb{R}} V \to \prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^N \times_{\mathbb{R}} \prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^N.$$

Consider the product of projections

$$(\pi,\pi) : (\mathbb{A}^{2N}_{\mathbb{R}} \setminus \{(0,\ldots,0)\}) \times_{\mathbb{R}} (\mathbb{A}^{2N}_{\mathbb{R}} \setminus \{(0,\ldots,0)\}) \to \prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^N \times_{\mathbb{R}} \prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^N.$$

Similarly to the construction of the join, we consider the Zariski closure of the image of $(\pi,\pi)^{-1}(W)$ under the addition morphism

$$\mathbb{A}^{2N}_{\mathbb{R}} \times_{\mathbb{R}} \mathbb{A}^{2N}_{\mathbb{R}} \to \mathbb{A}^{2N}_{\mathbb{R}} : (v_1, v_2) \mapsto v_1 + v_2,$$

remove the origin, and take the image under π. The result of these operations is a closed subvariety X of $\prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^N$ of dimension at most

$$4(n_1-1)+\cdots+4(n_d-1)-2.$$

By construction, the set of real points $X(\mathbb{R})$ corresponds to a subset of $\mathbb{PC}^N$ that contains all rank-2 tensors admitting a weak two-orthogonal decomposition.

Now we write r as $2r_0+\delta$ with r_0 a nonnegative integer and δ an element in $\{0,1\}$. We set

$$\mathcal{J} = J(\underbrace{X,\ldots,X}_{r_0 \text{ copies}}, V^{\delta}) \subset \prod_{\mathbb{C}/\mathbb{R}} \mathbb{PC}^N,$$

where the notation V^{δ} means that we include the variety V only in the case where $\delta = 1$. The join variety $\mathcal{J}$ has dimension at most

$$r_0 \mathrm{dim}(X) + \delta \mathrm{dim}(V) + 2(r_0+\delta-1),$$

which is bounded from above by

$$r_0\left(-2+4\sum_{\ell=1}^{d}(n_\ell-1)\right)+2\delta\sum_{\ell=1}^{d}(n_\ell-1)+2(r_0+\delta-1) = 2r\sum_{\ell=1}^{d}(n_\ell-1)+2\delta-2;$$

the set of real points $\mathcal{J}(\mathbb{R})$ corresponds to a subset of $\mathbb{PC}^N$ that contains the set $\sigma_r^{\perp}(\mathcal{S})$ of rank r tensors admitting a weak two-orthogonal decomposition. ∎

The established upper bound may be coarse, because the construction of X and $\mathcal{J}$ in the above proof only takes the weak two-orthogonality into account for the successive rank-1 terms $2k$ and $2k+1$ for all k, in a weak two-orthogonal decomposition. Nevertheless, this bound is sufficient for proving the main theorem.

With the understanding of the dimension of $\sigma_r^{\perp}$ in place, we can now state the following.

Lemma 4.2. *If*

$$\dim(\sigma_r^{\perp}) < \dim\left(\prod_{\mathbb{C}/\mathbb{R}} \sigma_r\right) = 2\dim(\sigma_r),$$

then the set of rank-r tensors not admitting a rank-r weak two-orthogonal decomposition, and, consequently, not admitting an SEY decomposition, i.e.,

$$\left(\prod_{\mathbb{C}/\mathbb{R}} \sigma_r \setminus \sigma_r^{\perp}\right)(\mathbb{R}),$$

is a Zariski-open subset, and, hence, dense open subset in the Euclidean topology, of

$$\left(\prod_{\mathbb{C}/\mathbb{R}} \sigma_r\right)(\mathbb{R}) = \sigma_r(\mathbb{C}).$$

If the above lemma applies, we will say that a *generic* rank-r tensor has no SEY decomposition. Beware that the term "generic" refers to the algebraic structure on the *real* variety $\prod_{\mathbb{C}/\mathbb{R}} \sigma_r$; this does not necessarily imply the existence of a Zariski-dense open subset of the *complex* variety σ_r whose points do not have an SEY decomposition.

For applying Lemma 4.2, we still need a lower bound on the dimension of σ_r. The dimensions of this variety have been studied for over a century now, but, unfortunately, they still elude the scientific community. The *expected dimension* of σ_r is well known,

$$\dim_E \sigma_r(\mathcal{S}) = \min\left\{N-1, (r-1) + r\sum_{\ell=1}^{d}(n_\ell - 1)\right\},$$

but in some instances $\dim \sigma_r(\mathcal{S})$ may be strictly smaller than the expected dimension; then, σ_r is called a *defective* rth order secant variety, and $\mathcal{S}$ a defective Segre variety. Only a limited number of defective secant varieties of Segre varieties are known, see, e.g., [1, section 6.1] and [31, section 5.5], while several secant varieties have been proved to be nondefective [1, 11, 12, 32, 45].[6] It is important to note that

$$\dim \sigma_r^{\perp} \le 2r\sum_{\ell=1}^{d}(n_\ell - 1) < 2\dim_E \sigma_r$$

whenever $d \ge 3$ and $r \ge 2$. That is, *whenever the* r*th order secant variety* σ_r *of a* d*-factor Segre variety is nondefective, a generic element of* σ_r *cannot admit a weak two-orthogonal decomposition of rank* r. This proves Theorem 1.4 from the introduction. Remark, further, that the discrepancy in dimension is $2(r-1)$, which provides ample leeway in the defectivity of σ_r before the approach outlined in this paper becomes moot. Combining the above observations with Lemma 4.2 and using the known results from the literature, we obtain the following.

COROLLARY 4.3. *Assume, without loss of generality, that* $2 \le n_1 \le n_2 \le \cdots \le n_d$ *with* $d \ge 3$. *Then, a generic*[7] *rank-*r *tensor in* $\mathbb{PC}^{n_1\times n_2\times\cdots\times n_d}$ *does not admit an SEY decomposition if*

$$2 \le r \le \max\left\{2, \left\lfloor \frac{n_1^d}{dn_1 - d + 1}\right\rfloor - n_1 + 1, \min\left\{n_d, \prod_{\ell=1}^{d-1} n_\ell - \sum_{\ell=1}^{d-1}(n_\ell - 1)\right\}\right\}.$$

Proof. The first item in the maximization is classic; see, e.g., [1]. From [1, Theorem 5.2] we know that $\sigma_r(\mathcal{S})$ is nondefective whenever r is smaller than the second item in the maximization. Then, using [1, Proposition 3.11], it follows that $\sigma_r(\mathcal{S})$ is nondefective because $n_i \ge n_1$. Applying Lemma 4.2 concludes this case. The third item in the maximization follows from combining Theorems 4.3 and 4.4 from [1], which summarize [11], with Lemma 4.2. ∎

Corollary 4.3 provides an easy-to-check condition on the rank r for which we know that admittance of an SEY decomposition is impossible for generic rank-r tensors in the tensor space of the stated dimensions. From the formula it follows that, for

[6] A defective rth order secant variety also implies that all secants of order $r' > r$ are defective, unless $\sigma_{r'}$ fills the ambient space $\mathbb{PC}^{n_1\times n_2\times\cdots\times n_d}$, so finding defective secant varieties should be an easy task [1]; yet both theoretical and probabilistic [49] results show that few varieties are defective.

[7] See the remark following Lemma 4.2.

sufficiently small r, an SEY decomposition cannot be admitted by a generic rank-r tensor. This condition is by no means necessary; it is a weak sufficient condition.[8]

The upper bound on the rank in Corollary 4.3 can often be improved by observing that the dimension of $\sigma_r^\perp$ is substantially smaller than σ_r, and that this discrepancy increases proportionally to r. We will illustrate this with one well-studied case: the cubic tensor spaces $\mathbb{P}\mathbb{C}^{n\times n\times\cdots\times n}$.

THEOREM 4.4. *Let $d \geq 3$, $n \geq 2$, $(d,n) \neq (3,3)$, and $(d,n) \neq (4,2)$. Then, a generic tensor in $\mathbb{C}^{n\times n\times\cdots\times n}$ (d times) does not admit an SEY decomposition.*

Proof. Let $\mathcal{S} = \mathcal{S}^{\mathbb{C}}_{n,n,\ldots,n}$, and let $\underline{r}$ denote the smallest $\underline{r}$ such that $\sigma_{\underline{r}} = \mathbb{P}\mathbb{C}^{n\times n\times\cdots\times n}$; this is called the *generic rank* [31]. A generic rank-$\underline{r}$ tensor is called a generic tensor. The *expected generic rank* for cubic tensor spaces is given by

$$\underline{r}_E = \left\lceil \frac{n^d}{dn - d + 1} \right\rceil \leq \underline{r}.$$

The generic rank equals the expected generic rank whenever the Segre variety $\mathcal{S}$ has no defective secant varieties. We define also the cumulative secant defect

$$\delta_r := \sum_{i=2}^{r} (\dim \sigma_{i-1} + \dim \mathcal{S} + 1 - \dim \sigma_i);$$

note that this quantity can only increase with r. Then, one verifies that

$$\dim \sigma_r = r \sum_{\ell=1}^{d} (n_\ell - 1) + (r-1) - \delta_r + 1.$$

It follows that Lemma 4.2 may be invoked whenever the middle inequality in

$$\frac{1}{2} \dim \sigma_r^\perp \leq r \sum_{\ell=1}^{d} (n_\ell - 1) + \delta - 1 < \dim \sigma_r = r \sum_{\ell=1}^{d} (n_\ell - 1) + (r-1) - \delta_r + 1$$

holds; note that this is not a *necessary* condition, however, because the upper bound in Proposition 4.1 is likely to be coarse. The middle inequality is then equivalent with

$$-\delta + 1 + (r-1) > \delta_r - 1, \text{ which is satisfied if } r > \delta_r + \delta - 1.$$

Recalling that $\delta \in \{0,1\}$, we find that $r > \delta_r$ implies the latter equation regardless of δ and is thus a sufficient condition for $\dim \sigma_r^\perp < 2 \dim \sigma_r$. The foregoing discussion is only valid when r is small enough so that $\sigma_r^\perp$ does not fill the ambient space, but this will not arise in the remainder.

If $d = 3$, Lickteig's classic result [32, p. 97] on the nondefectivity of $\mathcal{S}$, $n \neq 3$, can be used. It states that all secant varieties of $\mathcal{S}$ are nondefective, and, thus, $\underline{r} = \underline{r}_E$. Therefore, $\delta_r = 0$ for $r < \underline{r}_E$, and $\delta_{\underline{r}_E} < 3n - 3$. Thus, if the last inequality in

$$\underline{r}_E = \left\lceil \frac{n^3}{3n-2} \right\rceil \geq \frac{n^3}{3n-2} > 3n - 3 \tag{4.1}$$

[8]In fact, Abo, Ottaviani, and Peterson conjectured in [1] that, aside from the known exceptions, $\sigma_r(\mathcal{S})$ is nondefective if r is strictly less than the generic rank of the space $\mathbb{P}\mathbb{C}^{n_1\times\cdots\times n_d}$; this conjecture was proved for $r \leq 6$ in [1] and for $n_1 \cdots n_d \leq 100$ in [7, Theorem 7.5].

is satisfied, then $\dim \sigma_r^\perp < 2 \dim \sigma_r$, so that Lemma 4.2 applies. It is straightforward to verify using a computer algebra system that the last inequality in (4.1) is satisfied whenever $n \geq 7$. The remaining cases can be verified by substituting $n = 2, 4, 5, 6$ in the more refined sufficient condition $\underline{r}_E > \delta_{\underline{r}_E} = \underline{r}_E(3n-2) - n^3$. Note that this sufficient condition is *not* valid for the *defective* Segre variety $\mathbb{PC}^3 \times \mathbb{PC}^3 \times \mathbb{PC}^3$, and one may verify that $\underline{r}_E = 4 < \underline{r} = 5$ and $\underline{r} = 5 \not> \delta_{\underline{r}} = 8$.

If $d \geq 4$, it follows from [1, Theorem 5.2] that $\delta_r = 0$ for $r \leq \frac{n^d}{dn-d+1} - n$. In the same theorem it is stated that there are at most n defective secants of $\mathcal{S}$ before the space is filled. Consequently, the maximum cumulative secant defect in the generic rank is $n(dn-d+1)$. If we compare this with the last rank for which we know that the cumulative defect is zero, then $r > \delta_r$ for all r in that range. Therefore, if

$$\frac{n^d}{dn-d+1} - n > n(dn-d+1), \text{ i.e., } n^{d-1} > (dn-d+1)(dn-d+2), \tag{4.2}$$

it follows that $\dim \sigma_r^\perp < 2 \dim \sigma_r$, and Lemma 4.2 applies. If $d = 8$, one immediately obtains $n \geq 2$, providing the base case for the following inductive proof. Assume that

$$n^{k-1} > (kn-k+1)(kn-k+2)$$

for all $n \geq 2$, for some $k \geq 8$. Consider, then, the fraction

$$\begin{aligned}\alpha(k,n) &= \frac{((k+1)n-(k+1)+1)((k+1)n-(k+1)+2)}{(kn-k+1)(kn-k+2)} \\ &= \left(\frac{k+1}{k}\right)^2 - \frac{k-1}{k^2(kn-k+1)} - \frac{2(2+k)}{k^2(kn-k+2)}.\end{aligned}$$

It is easy to verify that $\alpha(k,n) < (9/8)^2 < 2$ for $n \geq 2$ and $k \geq 8$. Clearly, if $\alpha(k,n) < 2$, we have

$$\begin{aligned}n^k > \alpha(k,n)n^{k-1} &> \alpha(k,n)(kn-k+1)(kn-k+2) \\ &= ((k+1)n-(k+1)+1)((k+1)n-(k+1)+2),\end{aligned}$$

proving the inductive case. For $d = 6, 7$, one finds, e.g., using a computer algebra system, that (4.2) is satisfied if $n \geq 3$. Using Catalisano, Geramita, and Gimigliano's recent result on the nondefectivity of $n = 2$ if $d \geq 5$ [12], we may verify that the more refined inequality $\underline{r} = 6 > 4 = \delta_{\underline{r}} = 6\underline{r} - 2^5$ holds for $d = 5$, that $\underline{r} = 10 > 6 = \delta_{\underline{r}}$ for $d = 6$, and that $\underline{r} = 16 > 0 = \delta_{\underline{r}}$ for $d = 7$. One can also verify that (4.2) holds if $n \geq 15$ for $d = 4$ and if $n \geq 5$ for $d = 5$. For proving the remaining cases, we invoke the next lemma.

LEMMA 4.5. *The Segre variety $\mathbb{PC}^n \times \cdots \times \mathbb{PC}^n$ (d times) is nondefective for $d = 4$ and $3 \leq n \leq 14$, and for $d = 5$ and $2 \leq n \leq 4$.*

Proof. Gesmundo [21] already proved the cases $n = 3, \ldots, 8$ and $n = 10$ for $d = 4$. The remaining cases were proved by a computer program that exploits Terracini's lemma [46] in the classic manner to construct the tangent space to the rth order secant variety of a Segre in a generic point, and computes the rank of this matrix representation using Gaussian elimination, which corresponds to the dimension of the secant variety; see, e.g., [49, section 1]. Exploiting an observation by G. Ottaviani, all computations were performed in a prime finite field of size $2^{13} - 1$: whenever the tangent space matrix is of maximal rank in such a field, it also has the maximal rank in $\mathbb{C}$. Using this algorithm, we verified that the dimension of the secant varieties was as expected. $\square$

By straightforward calculations it follows that $\underline{r}_E = \underline{r}$, by the lemma, is strictly larger than $\delta_{\underline{r}} = \underline{r}(dn - d + 1) - n^d$ for each of the remaining cases. $\square$

Finally, we prove the main result, already presented in the introduction as Theorem 1.3. It states that in every complex tensor space of order at least three, one can always find a set of positive Lebesgue measure wherein its elements do not admit an SEY decomposition. The previous theorem actually specializes the main theorem for cubic tensor spaces, proving that the set is then dense.

THEOREM 4.6. *Let $d \geq 3$. Then, there exists a nonempty open subset V of $\mathbb{C}^{n_1 \cdots n_d} \cong \mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$ with respect to the Euclidean topology such that the points in V do not admit an SEY decomposition; V has positive Lebesgue measure.*

Proof. Set $N = n_1 \cdots n_d$. It is known that $\sigma_2(\mathcal{S})$ is never defective for $d \geq 3$ [1, p. 781]. Let $\mathcal{A}_0$ be a rank-2 tensor in $\mathbb{C}^N$ that does not lie in the affine cone over $\sigma_2^\perp \subset \mathbb{P}\mathbb{C}^N$. Since $\sigma_2^\perp$ is closed in $\mathbb{P}\mathbb{C}^N$ with respect to the Euclidean topology, there exists a real value $\epsilon > 0$ such that the open ℓ_2-ball around $\mathcal{A}_0$ in $\mathbb{C}^N$ with radius ϵ, i.e., the set of points at distance at most ϵ from $\mathcal{A}_0$ as measured in the Euclidean norm, is disjoint from the cone over $\sigma_2^\perp$. Now let V be the open ℓ_2-ball around $\mathcal{A}_0$ in $\mathbb{C}^N$ with radius $\epsilon/2$. Then for every point $\mathcal{A}$ in V, any best rank-2 approximation $\mathcal{A}^*$ of $\mathcal{A}$ satisfies

$$\|\mathcal{A} - \mathcal{A}^*\| \leq \|\mathcal{A} - \mathcal{A}_0\| < \frac{\epsilon}{2} \quad \text{so that} \quad \|\mathcal{A}^* - \mathcal{A}_0\| < \epsilon,$$

and, hence, $\mathcal{A}^*$ does not lie in the affine cone over $\sigma_2^\perp$. This implies that $\mathcal{A}^*$ does not admit an SEY decomposition. Alternatively, if $\mathcal{A}$ does not have a best rank-2 approximation, we again find that it does not admit an SEY decomposition. In both cases, $\mathcal{A}$ does not admit a best rank-2 approximation that itself admits an SEY decomposition. However, from the definition it follows that a tensor $\mathcal{A}$ can only admit an SEY decomposition if for *every rank* there exists a best approximation of that rank that itself admits an SEY decomposition; this concludes the proof. $\square$

Note that in the proof we may substitute σ_2 and $\sigma_2^\perp$ for any other secant variety that would satisfy Lemma 4.2, but such an exercise would only be useful if this would somehow provide information about the value of ϵ, and possibly increase it.

5. Conclusions. We argued that current approaches for investigating the existence of an SEY decomposition rely explicitly on the existence of a set of tensors admitting supergeneric ranks with positive measure. Such an approach fails in a complex setting, leading us to propose an alternative strategy based on algebraic geometry and a comparison of dimensions of the varieties involved. We showed, for every complex tensor space, that an SEY decomposition is not admitted at least by a set of positive Lebesgue measure. Notwithstanding these results, we also provided a nontrivial class of tensors, i.e., those admitting a strong two-orthogonal decomposition, which are optimally truncatable.

Acknowledgments. J. Landsberg is heartily thanked for introducing us to algebraic geometry by means of his nice talk at the 17th International Linear Algebra Society Conference in Braunschweig, Germany, and for valuable insights and suggestions he provided there and in subsequent e-mail communication. G. Ottaviani is kindly thanked for the feedback he provided while the first author visited the Università di Firenze. We thank F. Cools for identifying an error in the proof of Theorem 3.3. Finally, C. Mertens is kindly thanked for various enlightening discussions we had.

We thank the referees for their detailed reading and helpful suggestions that improved this manuscript.

REFERENCES

[1] H. Abo, G. Ottaviani, and C. Peterson, *Induction for secant varieties of Segre varieties*, Trans. Amer. Math. Soc., 361 (2009), pp. 767–792.

[2] C. Appellof and E. Davidson, *Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents*, Anal. Chem., 53 (1981), pp. 2053–2056.

[3] M. Bennani Dosse, J. ten Berge, and J. Tendeiro, *Some new results on orthogonally constrained candecomp*, J. Classification, 28 (2011), pp. 144–155.

[4] D. Bini, *Border rank of a* $p \times q \times 2$ *tensor and the optimal approximation of a pair of bilinear forms*, in Automata, Languages and Programming, J. de Bakker and J. van Leeuwen, eds., Lecture Notes in Comput. Sci. 85, Springer, New York, 1980, pp. 98–108.

[5] D. Bini, *Relations between exact and approximate bilinear algorithms*, Calcolo, 17 (1980), pp. 87–97.

[6] D. Bini, G. Lotti, and F. Romani, *Approximate solutions for the bilinear form computational problem*, SIAM J. Comput., 9 (1980), pp. 692–697.

[7] C. Bocci, L. Chiantini, and G. Ottaviani, *Refined methods for the identifiability of tensors*, Ann. Mat. Pura. Appl. (2013).

[8] M. Brambilla and G. Ottaviani, *On the Alexander–Hirschowitz theorem*, J. Pure Appl. Algebra, 212 (2008), pp. 1229–1251.

[9] J. Cardoso and A. Souloumiac, *Blind beamforming for non-Gaussian signals*, IEEE Proceedings-F, 140 (1993), pp. 362–370.

[10] J. Carroll and J.-J. Chang, *Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart–Young" decomposition*, Psychometrika, 35 (1970), pp. 283–319.

[11] M. Catalisano, A. Geramita, and A. Gimigliano, *Ranks of tensors, secant varieties of Segre varieties and fat points*, Linear Algebra Appl., 355 (2002), pp. 263–285.

[12] M. Catalisano, A. Geramita, and A. Gimigliano, *Secant varieties of* $\mathbb{P}^1 \times \cdots \times \mathbb{P}^1$ *(n-times) are not defective for* $n \geq 5$, J. Algebraic Geom., 20 (2011), pp. 295–327.

[13] J. Chen and Y. Saad, *On the tensor SVD and the optimal low rank orthogonal approximation of tensors*, SIAM J. Matrix Anal. Appl., 30 (2009), pp. 1709–1734.

[14] L. Chiantini and G. Ottaviani, *On generic identifiability of* 3*-tensors of small rank*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 1018–1037.

[15] P. Comon, G. H. Golub, L.-H. Lim, and B. Mourrain, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279.

[16] L. De Lathauwer, *Decompositions of a higher-order tensor in block terms—part* II: *definitions and uniqueness*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1033–1066.

[17] L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.

[18] L. De Lathauwer and A. Stegeman, *A method to avoid diverging components in the Candecomp/Parafac model for generic* $I \times J \times 2$ *arrays*, SIAM J. Matrix Anal. Appl., 30 (2009), pp. 1614–1638.

[19] V. de Silva and L.-H. Lim, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1084–1127.

[20] C. Eckart and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika, 1 (1936), pp. 211–218.

[21] F. Gesmundo, *An asymptotic bound for secant varieties of Segre varieties*, Ann. Univ. Ferrara, 59 (2013), pp. 285–302.

[22] L. Grasedyck, *Hierarchical singular value decomposition of tensors*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2029–2054.

[23] W. Hackbusch, *Tensor Spaces and Numerical Tensor Calcalus*, Springer Ser. Comput. Math. 42, Springer, Berlin, 2012.

[24] W. Hackbusch and S. Kühn, *A new scheme for the tensor representation*, J. Fourier Anal. Appl., 15 (2009), pp. 706–722.

[25] R. A. Harshman, *Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis*, UCLA Working Papers in Phonetics, 16 (1970), pp. 1–84.

[26] F. Hitchcock, *The expression of a tensor or a polyadic as a sum of products*, J. Math. Phys., 6 (1927), pp. 164–189.

[27] F. Hitchcock, *Multiple invariants and generalized rank of a p-way matrix or tensor*, J. Math. Phys., 7 (1927), pp. 39–79.

[28] T. G. Kolda, *Orthogonal tensor decompositions*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 243–255.

[29] T. G. Kolda, *A counterexample to the possibility of an extension of the Eckart–Young low-rank approximation theorem for the orthogonal rank tensor decomposition*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 762–767.

[30] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.

[31] J. M. Landsberg, *Tensors: Geometry and Applications*, Grad. Stud. Math. 128, AMS, Providence, RI, 2012.

[32] T. Lickteig, *Typical tensorial rank*, Linear Algebra Appl., 69 (1985), pp. 95–120.

[33] M. Mørup, *Applications of tensor (multiway array) factorizations and decompositions in data mining*, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1 (2011), pp. 24–40.

[34] I. Oseledets, *Tensor-train decomposition*, SIAM J. Sci. Comput., 33 (2011), pp. 2295–2317.

[35] P. Paatero, *Construction and analysis of degenerate PARAFAC models*, J. Chemometrics, 14 (2000), pp. 285–299.

[36] B. Pesquet-Popescu, J.-C. Pesquet, and A. Petropulu, *Joint singular value decomposition—a new tool for separable representation of images*, in Proceedings of the 2001 International Conference on Image Processing, 2001, pp. 569–572.

[37] E. Schmidt, *Zur Theorie der linearen und nichtlinearen Integralgleichungen*, Math. Ann., 63 (1907), pp. 433–476.

[38] A. Smilde, R. Bro, and P. Geladi, *Multiway Analysis: Applications in the Chemical Sciences*, John Wiley & Sons, Chichester, England, 2004.

[39] A. Stegeman, *Degeneracy in CANDECOMP/PARAFAC explained for $p \times p \times 2$ arrays of rank $p+1$ or higher*, Psychometrika, 71 (2006), pp. 483–501.

[40] A. Stegeman, *Degeneracy in CANDECOMP/PARAFAC and INDSCAL explained for several three-sliced arrays with a two-valued typical rank*, Psychometrika, 72 (2007), pp. 601–619.

[41] A. Stegeman, *Low-rank approximation of generic $p \times q \times 2$ arrays and diverging components in the CANDECOMP/PARAFAC model*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 988–1007.

[42] A. Stegeman, *Candecomp/Parafac: from diverging components to a decomposition in block terms*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 291–316.

[43] A. Stegeman, *A three-way Jordan canonical form as limit of low-rank tensor approximations*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 624–650.

[44] G. W. Stewart, *On the early history of the singular value decomposition*, SIAM Rev., 35 (1993), pp. 551–566.

[45] V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra Appl., 52–53 (1983), pp. 645–685.

[46] A. Terracini, *Sulla V_k per cui la varietà degli S_h $h+1$-secanti ha dimensione minore dell'ordinario*, Rend. Circ. Mat. Palermo, 31 (1911), pp. 392–396.

[47] L. R. Tucker, *Some mathematical notes on three-mode factor analysis*, Psychometrika, 31 (1966), pp. 279–311.

[48] N. Vannieuwenhoven, R. Vandebril, and K. Meerbergen, *A new truncation strategy for the higher-order singular value decomposition*, SIAM J. Sci. Comput., 34 (2012), pp. A1027–A1052.

[49] N. Vannieuwenhoven, R. Vandebril, and K. Meerbergen, *A randomized algorithm for testing nonsingularity of structured matrices with an application to asserting nondefectivity of segre varieties*, IMA J. Numer. Anal., 2014 (2014), pp. 1–34.

[50] T. Zhang and G. H. Golub, *Rank-one approximation to high order tensors*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 534–550.